Two-Layer Mutually Reinforced Random Walk for Improved Multi-Party Meeting Summarization

Yun-Nung (Vivian) Chen yvchen@cs.cmu.edu

LTI Student Research Symposium – Aug. 22nd, 2012 Language Technologies Institute · Carnegie Mellon University



Outline





Outline Introduction Approach Experiments Conclusion

Motivation

Extractive Summarization



Outline Introduction Approach Experiments Conclusion

Motivation

Extractive Summarization

Motivation

- Ø Speech Summarization
 - Spoken documents are more difficult to browse than texts
 - \rightarrow easy to browse, save time, easily get the key points
- Multi-Party Corpus
 - Speaker information may help summarization



Outline Introduction Approach Experiments Conclusion

Motivation

Æ Extractive Summarization

Extractive Summarization (1/2)

- *o* Extractive Speech Summarization
 - Select the indicative utterances in a spoken document
 - Cascade the utterances to form a summary



Extractive Summarization (2/2)

- Selection of Indicative Utterances
 - Each utterance U in a spoken document d is given an importance score I(U, d)
 - Select the indicative utterances based on I(U,d)
 - The number of utterances selected as summary is decided by a predefined ratio





Ø Graph Construction

- Two-Layer Mutually Reinforced Random Walk
 - Ø Between-Layer Propagation (MRRW-BP)
 - Ø Within- and Between-Layer Propagation (MRRW-WBP)



Ø Graph Construction

- Two-Layer Mutually Reinforced Random Walk
 - Ø Between-Layer Propagation (MRRW-BP)
 - Within- and Between-Layer Propagation (MRRW-WBP)

Graph Construction (1/2)

- OUtterance-Layer
 - *i* Each node is the utterance in the meeting document
 - The edge is weighted by topical/lexical similarity between two utterances



Graph Construction (2/2)

Ø Speaker-Layer

- Each node is the speaker in the meeting document
 - Combine all utterances from the same speaker as the speaker node
- The edge (red and green) is weighted by lexical similarity between two nodes





Ø Graph Construction

Two-Layer Mutually Reinforced Random Walk

- Ø Between-Layer Propagation (MRRW-BP)
- Within- and Between-Layer Propagation (MRRW-WBP)

Two-Layer Mutually Reinforced Random Walk (1/3)

Similarity Matrix

- LUU: utterance-to-utterance relation (topical/lexical similarity)
- L_{ss}: speaker-to-speaker relation (TF-IDF cosine similarity)
- L_{US}: utterance-to-speaker relation (TF-IDF cosine similarity)
- L_{SU}: speaker-to-utterance relation (TF-IDF cosine similarity)



Two-Layer Mutually Reinforced Random Walk (1/3)

Similarity Matrix

- LUU: utterance-to-utterance relation (topical/lexical similarity)
- L_{ss}: speaker-to-speaker relation (TF-IDF cosine similarity)
- L_{US}: utterance-to-speaker relation (TF-IDF cosine similarity)
- L_{SU}: speaker-to-utterance relation (TF-IDF cosine similarity)



Two-Layer Mutually Reinforced Random Walk (2/3)

O Topical similarity between utterances

✓ Edge weight TopicSim(U_i, U_i) (utterance U_i → utterance U_i)



TopicSim(U_i, U_i): latent topic similarities of U_i to U_i based on PLSA model

Two-Layer Mutually Reinforced Random Walk (3/3)

Lexical similarity between utterances

Cosine similarity between TF-IDF vectors from U_i and U_i

LexSim(U_i, U_i): evaluated by the word overlap between two utterances



Ø Graph Construction

- Two-Layer Mutually Reinforced Random Walk
- Ø Between-Layer Propagation (MRRW-BP)
 - Within- and Between-Layer Propagation (MRRW-WBP)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$



$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$



original importance of utterance

$$\begin{cases}
F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\
F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)}
\end{cases}$$



Mathematical Formulation

scores propagated from speaker-layer

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$



Mathematical Formulation

scores propagated from speaker-layer

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$







$$\begin{cases} F_U^{(t+1)} = (1 - \alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1 - \alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \\ \sigma \text{iginal importance of speaker} \end{cases}$$

• Original importance
• Utterance
 $I(U, d) \Rightarrow \text{baseline}$
• Speaker
Equal weight
Utterance-Layer

Mathematical Formulation

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$

scores propagated from utterance-layer



Mathematical Formulation

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$

scores propagated from utterance-layer



Mathematical Formulation

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$

Utterance node U can get higher score when

- ① Higher original importance I(U, d)
- ② More speaker nodes similar to utterance U

Mathematical Formulation

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SU}F_U^{(t)} \end{cases}$$

Converged Solution

$$\begin{cases} F_{U}^{*} = (1 - \alpha) F_{U}^{(0)} + \alpha \cdot L_{US} F_{S}^{*} \\ F_{S}^{*} = (1 - \alpha) F_{S}^{(0)} + \alpha \cdot L_{SU} F_{U}^{*} \end{cases}$$

$$F_{U}^{*} = (1 - \alpha) F_{U}^{(0)} + \alpha \cdot L_{US} \left((1 - \alpha) F_{S}^{(0)} + \alpha \cdot L_{SU} F_{U}^{*} \right)$$

$$= \left((1 - \alpha) F_{U}^{(0)} e^{T} + \alpha (1 - \alpha) L_{US} F_{S}^{(0)} e^{T} + \alpha^{2} L_{US} L_{SU} \right) F_{U}^{*}$$

$$= M_{1} F_{U}^{*}$$

 F_{U}^{*} (final utterance scores) is the dominate eigenvector of M_{1}



Ø Graph Construction

- Two-Layer Mutually Reinforced Random Walk
 - Ø Between-Layer Propagation (MRRW-BP)
- > Ø Within- and Between-Layer Propagation (MRRW-WBP)

(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_{UU}^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$



(Within- and Between-Layer Propagation) (1/2)

Mathematical Formulation

$$\begin{cases} F_U^{(t+1)} = (1-\alpha)F_U^{(0)} + \alpha \cdot L_U^T L_{US}F_S^{(t)} \\ F_S^{(t+1)} = (1-\alpha)F_S^{(0)} + \alpha \cdot L_{SS}^T L_{SU}F_U^{(t)} \end{cases}$$

Utterance node U can get higher score when

- ① Higher original importance I(U, d)
- ② More speaker nodes similar to utterance U
- ③ More important utterances similar to utterance U

Converged Solution

$$F_{U}^{*} = \left((1-\alpha)F_{U}^{(0)}e^{T} + \alpha(1-\alpha)L_{UU}^{T}L_{US}F_{S}^{(0)}e^{T} + \alpha^{2}L_{UU}^{T}L_{US}L_{SS}^{T}L_{SU} \right)F_{U}^{*}$$

= $M_{2}F_{U}^{*}$

 F_{U}^{*} (final utterance scores) is the dominate eigenvector of M_{2}



- *O* Experimental Setup
- Evaluation Metrics
- Results

Outline Introduction Approach Experiments Conclusion

- - Evaluation Metrics
 - Results

Experimental Setup

- CMU Speech Meeting Corpus
 - ✓ 10 meetings from 2006/04 2006/06
 - #Speaker: 6 (total), 2-4 (each meeting)
 - ✓ WER = 44%
- Reference Summaries
 - Manually labeled by two annotators
- Parameter Setting
 - ο α = 0.9
 - Extractive summary ratio = 30%

$$\begin{cases} F_{U}^{(t+1)} = (1-\alpha)F_{U}^{(0)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)F_{S}^{(0)} + \alpha \\ F_{U}^{(t+1)} = (1-\alpha)F_{U}^{(0)} + \alpha \\ F_{U}^{(t+1)} = (1-\alpha)F_{U}^{(0)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)F_{S}^{(0)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)F_{S}^{(0)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)F_{S}^{(0)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)F_{S}^{(t)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)F_{S}^{(t+1)} + \alpha \\ F_{S}^{(t+1)} = (1-\alpha)$$



- *O* Experimental Setup
- O Evaluation Metrics
 - Results

Evaluation Metrics

- ROUGE
 - ROUGE-1
 - F-measure of <u>matched unigram</u> between extracted summary and reference summary
 - ROUGE-L (Longest Common Subsequence)
 - F-measure of <u>matched LCS</u> between extracted summary and reference summary



- Experimental Setup
- Evaluation Metrics
- 🔪 🥖 Results



Results – Baseline & Single-Layer Graph Approaches



Graph-based approaches are significantly better than baseline

Results – Baseline & Single-Layer Graph Approaches



Improvement for ASR is larger than manual transcripts due to recognition errors



Two-layer approaches involving speaker info. outperform single-layer approaches



The utterances from the speakers who speak more important utterances tend to be more important

Results – Effectiveness of Within-Layer Propagation **ROUGE-L (ASR) ROUGE-1 (ASR)** $\frac{1}{2}$ 50.5 50 49.5 49 48.5 48 47.5 47 46.5 46 Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Layer Two-Laver Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Laver Two-Laver MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) (LexSim) (TopicSim) (LexSim) (TopicSim) **ROUGE-L** (Manual) **ROUGE-1** (Manual) 49 48.5 48 47.5 47 46.5 46 45.5 45 44.5 44 Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Laver Two-Laver Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Laver Two-Laver (LexSim) (TopicSim) MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) (LexSim) (TopicSim)

For ASR transcripts, within-layer propagation using topical similarity is better

(LexSim)

(TopicSim)

Results – Effectiveness of Within-Layer Propagation **ROUGE-L (ASR) ROUGE-1 (ASR)** $\frac{1}{2}$ 50.5 50 49.5 49 48.5 48 47.5 47 46.5 46 Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Layer Two-Laver Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Laver Two-Laver MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) (LexSim) (TopicSim) (LexSim) (TopicSim) **ROUGE-L** (Manual) **ROUGE-1** (Manual) 49 48.5 $\frac{1}{2}$ 48 47.5 47 46.5 46 45.5 45 44.5 44 Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Laver Two-Laver Baseline: LTE RandomWalk RandomWalk Two-Layer Two-Laver Two-Laver (TopicSim) (LexSim) MRRW-BP MRRW-WBP MRRW-WBP (LexSim) (TopicSim) MRRW-BP MRRW-WBP MRRW-WBP

For manual transcripts, within-layer propagation using lexical similarity is better

(TopicSim)

(LexSim)

Results – Lexical & Topical Similarity



For ASR transcripts, topical similarity outperforms lexical similarity in both cases

Results – Lexical & Topical Similarity



Topical similarity can compensate recognition errors

Results – Lexical & Topical Similarity



Lexical similarity from word overlap may have some noises due to recognition errors

Results – Lexical & Topical Similarity



For manual transcripts, lexical similarity outperforms topical similarity in both cases

Results – Lexical & Topical Similarity



Lexical similarity can model relations accurately since in absence of errors



Proposed approaches achieve 7.2% and 8.2% relative improvement compared to baseline for ASR and manual respectively



- Graph-based approaches can improve speech summarization performance
- Two-layer approaches involving speaker information can get further improvement
- Topical similarity is more robust to recognition errors
 - \rightarrow better for ASR transcripts
- ✓ Lexical similarity is more accurate when absence of errors
 → better for manual transcripts
- Our proposed approaches achieve more than 7% relative improvement compared to the baseline

Thanks for your attention! © Q&A